Predicting Alignment Risk to Prevent Localization Failure

Predicting Alignment Risk to Prevent Localization Failure

Simona Nobili[1,2], Georgi Tinchev[2] and Maurice Fallon[2]

***Abstract*— During localization and mapping the success of point cloud registration can be compromised when there is an absence of geometric features or constraints in corridors or across doorways, or when the volumes scanned only partly overlap, due to occlusions or constrictions between subsequent observations. This work proposes a strategy to predict and prevent laser-based localization failure. Our solution relies on explicit analysis of the point cloud content prior to registration. A model predicting the risk of a failed alignment is learned by analysing the degree of spatial overlap between two input point clouds and the geometric constraints available within the region of overlap. We define a novel measure of alignability for these constraints. The method is evaluated against three real-world datasets and compared to baseline approaches. The experiments demonstrate how our approach can help improve the reliability of laser-based localization during exploration of unknown and cluttered man-made environments.**

I. Introduction

Simultaneous Localization and Mapping (SLAM) systems need to be able to reliably operate with a low failure rate for long periods of time and in a variety of environments. These systems should include fail-safe operation modes and auto-tuning capabilities so as to adapt to different challenges and achieve the vision of robot perception described in [1].

The work presented in this paper is focused on guaranteeing fail-safe operation of laser-based localization systems during exploration of cluttered man-made environments.

State-of-the-art laser-based localization systems have been demonstrated to achieve low drift over long distances but can also be easily induced to fail in many real-world scenarios. A cause of failure is the absence of geometric features which are necessary to constrain the alignment between two point clouds. For example, long corridors are unconstrained because of missing geometric features in one dimension. Failures can also occur when passing through doorways, or due to occlusions which cause large variations in the volume scanned by consecutive sensor sweeps. We define *alignability* as a measure of the capacity for two point clouds to be aligned given their geometric constraints, e.g., mutually visible planar surfaces.

In many cases sensors have a limited field-of-view (FOV) because they are physically integrated within a vehicle's chassis. With an obscured FOV, the degree of overlap between consecutive scans made by the sensor is gradually reduced. Occlusions and constrictions in cluttered environments also introduce a large degree of overlap variation. This is a major problem for point cloud registration algorithms, as they are sensitive to the degree of overlap between input clouds. We believe that there is a need for failure prediction which is reliable to varying overlap.

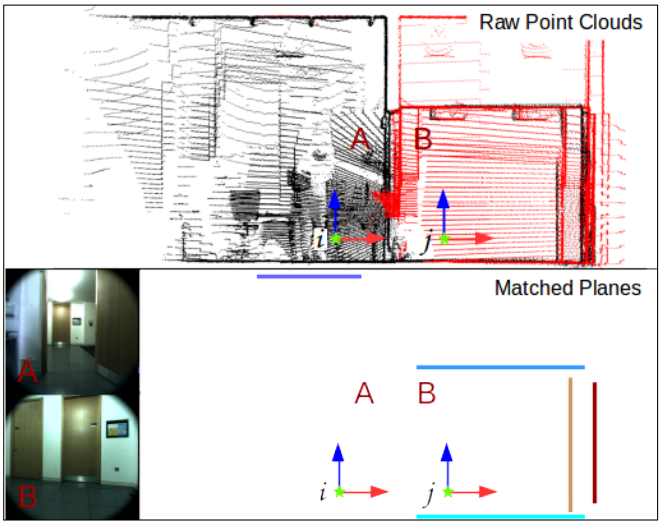

Fig. 1: The side view of a robot crossing the doorway between rooms A and B while capturing point clouds in poses i and j. According to our metric the clouds overlap by as little as 15%, and alignabilty is low due to the absence of matched planes in the direction of the reader. *Top*: The raw point clouds from each location. *Bottom*: Our proposed method analyses overlap and alignability, the latter derived when matching planes common to each point clouds. These are used to infer the suitability for alignment.

The main contribution of this work is a method which predicts the *risk of failed alignment* between two point clouds, which is learned as a function of the overlap between a reference and a reading cloud[1] and the constraints available in the region of overlap between them. An example of unconstrained alignment is illustrated in Fig. 1. In contrast to previous works, [3], [4], [5], our method accounts for the degree of overlap between point clouds. It is also independent of point-wise data association and the registration approach used. The pipeline of our approach is shown in Fig. 2, which assumes that an initial (drifting) estimate of the sensor's pose is available to initialize the alignment.

Our contribution is broken down as follows:

i) we define a novel overlap metric for 3D point clouds which takes into account the relative poses from which the clouds were captured, the structural features of the clouds, as well as the free space information. The metric gives improved performance with respect to previous work, particularly when occlusions occur,

ii) we derive an *alignability* metric which quantifies the

[1]S. Nobili is with the Institute of Perception, Action and Behaviour, School of Informatics, University of Edinburgh, UK.

[2]All authors are with the Oxford Robotics Institute, University of Oxford, UK. {snobili, gtinchev, mfallon}@robots.ox.ac.uk

[1]Using the notation from [2] we refer to the inputs to registration as a *reference* and a *reading* cloud with the latter to be aligned to the former.

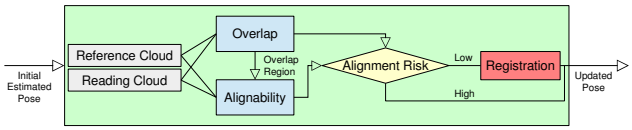

Fig. 2: The proposed pipeline.

degree to which alignment is constrained by exploiting the planar geometry commonly found in indoor environments. Our data-driven approach achieves a higher accuracy with respect to previous approaches,

iii) we learn a model which can predict *alignment risk*, based on these overlap and alignability measures. The model allows us to prevent registration failure when the geometry is unstable and the overlap is not uniform.

The remainder of the paper is structured as follows: Sec. II presents related approaches from the literature, Sec. III describes our algorithm in detail, Sec. IV presents an extensive experimental evaluation.

II. Related Work

Iterative Closest Point (ICP) is a very commonly used method for 3D point cloud alignment. Its basic formulation [6] estimates the relative alignment between two point clouds iteratively, through four main steps: data filtering, data association, outlier rejection and minimization of a point-to-point distance function. Several improvements to the algorithm have been proposed over the past years which focus on changing the minimized distance function to suit different environments [7], [8], [9].

However, the optimal registration estimated by ICP is not always a good solution. This can be because of an unconstrained degree of freedom such as when crossing a doorway (Fig. 1), or trying to align within a long corridor.

Previous works [3], [4], [5] have stated that the stability of registration can be evaluated after the set of point matches have been selected, by Principal Component Analysis (PCA) on the covariance matrix used for error minimization. If the covariance matrix is not full rank the registration is underconstrained. However, this analysis depends on the data association step of the registration algorithm. When the overlap is low, the number of point matches available for data association might be insufficient for the eigenvectors of the covariance matrix to stabilize, thus leading to an unreliable measure of the constraints. Motivated by this, in our work we focus on analysis **before** the point-wise data association step of the ICP algorithm, and we explicitly account for overlap variation.

Zhen *et al.* [10] formulated a *localizability* measure for a sensor with respect to a prior 3D map of a structure of interest. It was computed offline by generating synthetic laser data to simulate observations from within the map. They used it to plan trajectories of a UAV so as to stay within areas with high localizability. Having a different application in mind, in our work we extend their formulation to a measure of *alignability* between two subsequent clouds. This is computed online to account for the current dynamics of the scene.

Pathak *et al.* [11] recognized the sensitivity of registration to low data overlap and formulated two overlap metrics which can be used to study the reason of alignment failures. The first metric used octrees to model **occupied** regions and then derive an overlapping surface area between two point clouds. The second measured the number of pixels on range-images mutually visible from both scans. The metrics were computed given the ground-truth alignment between the clouds. Herein, we use octrees to model **occupied** and **free** space and then derive an overlapping volume between two point clouds. This makes the estimate more robust to initial misalignment.

Pomerleau *et al.* [2] formulated an overlap measure using the ratio of points from a first cloud for which there was a corresponding point in the second cloud. This approach is asymmetric, for example, when the reference cloud is larger than the reading cloud. In contrast, we define overlap so as to be resilient to this asymmetry.

Finally, in our previous work [12] we proposed a strategy for non-incremental 3D scene registration, called Auto-tuned ICP (AICP). AICP extended a baseline ICP implementation [2] to more reliably register point clouds with reduced overlap by automatically tuning an outlier-rejection filter to account for the degree of overlap of the sensor's footprint. This framework allowed accurate registration to a single reference point cloud despite significant motion by the robot. In this work we extend AICP with a more robust overlap parameter which corrects for some issues highlighted in Sec. IV.

In summary, our work differs from the state-of-the-art in that we learn a model for predicting alignment risk which is based on both the overlap and alignability parameters. We demonstrate the utility of our model in a general context, where no prior map is available and overlap between the input point clouds varies dramatically.

In a SLAM context, we address the problem of preventing failures from within the front-end module of the system, whereas other works focus on graph optimization strategies to remove outliers at a back-end level [13], [1].

III. Predicting Alignment Risk

In this section, we derive a continuous variable quantifying the risk of alignment failure when registering two point clouds. We first define measures for overlap and alignability of two point clouds. We then generate a meta-parameter which can be used to predict the risk of alignment failure.

A. Measuring Overlap

We define the overlap, $\Omega \in [0, 1]$, between two point clouds using the initial estimated alignment between them (from odometry for example), their structural features and information about free space which is directly available given the sensor's origin.

The reference cloud ${}^{w}C_i$ and the reading cloud ${}^{w}C_j$ are captured from two sensor poses i and j, expressed in the world coordinate frame w.

Using an octree structure [14], two corresponding octrees are constructed, ${}^{w}O_i$ and ${}^{w}O_j$. Each explicitly models both free and occupied space. From ${}^{w}O_i$ and ${}^{w}O_j$, another octree containing the set of common voxels is constructed, O^{ij}, which defines the volume of overlap between the clouds.

We define the overlap parameter as

$$\Omega = \min\left(\frac{|O^{ij}|}{|O_i|}, \frac{|O^{ij}|}{|O_j|}\right) \tag{1}$$

where $|\bullet|$ indicates the cardinality of voxels in an octree. Our volumetric representation of overlap is shown in Fig. 4 (right).

B. Measuring Alignability

We define *alignability* α as a measure of the geometric constraints which can be used to constrain alignment between a reference and a reading cloud. The alignment between a pair of 3D point clouds is well-constrained if the transform aligning them is constrained in all three dimensions. Intuitively, we envisage that at least three mutually visible non-parallel planes should exist.

In the following we describe the two steps of our strategy: (i) firstly, we match planes common to the input clouds and compute a matrix $\mathbf{N}$ as the set of normal directions extracted from these planes, (ii) secondly, we show how $\mathbf{N}$ represents the set of available constraints between the clouds. We compute α from PCA on the row vectors of $\mathbf{N}$.

Matching Plane Patches: Having segmented the reference point cloud into a set of plane patches[2], we select only the ones which belong to the volume of overlap between the clouds, O^{ij}. We consider the u$-th$ patch P_u from this set. Similarly, we define P_v from the set of patches extracted from the reading cloud.

For each pair P_u and P_v, we compute a matching score Ω_p which is defined as the degree of spatial overlap between the patches. We consider each patch as a set of points contained within a bounding box in $\mathbb{R}^3$. We define the bounding boxes as V_u and V_v.

Ω_p is computed from the set of points P_u^{ij} and P_v^{ij} belonging to the patches P_u and P_v and living in the volume of intersection between V_u and V_v, as shown in Fig. 3,

$$\Omega_p = \frac{|P_u^{ij}|}{|P_u|} \cdot \frac{|P_v^{ij}|}{|P_v|}. \tag{2}$$

In Eq. (2) $|\bullet|$ indicates the cardinality of a set.

The best match for P_v between any plane P_u is the one maximizing the overlap Ω_p. For a match to be accepted, both a condition on Ω_p and on the maximum angle between the normals must hold. The normal directions $\mathbf{n}_k^T = [n_{k_x}, n_{k_y}, n_{k_z}]$ are extracted per point $k \in [1 : M]$ (where

[2]We refer to plane patches as locally planar distributions of points. We adopt a region growing strategy for plane segmentation [15]. A patch is accepted only if it satisfies criteria about its planarity and dimensions (e.g., larger than 0.30 $\times$ 0.30 m).

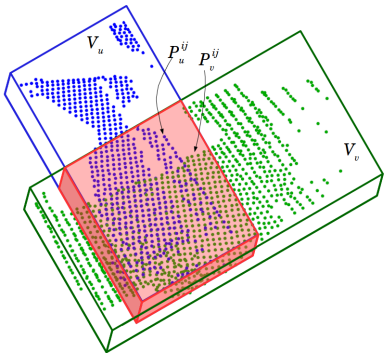

Fig. 3: Illustration of the plane matching strategy. A pair of plane patches P_u (blue points) and P_v (green points) are shown within their bounding boxes, V_u and V_v. The volume of intersection is shown in red. A score Ω_p for the match is computed from the percentage of points falling within this volume.

M is the number of points from all matched patches in the reading cloud), and $\mathbf{N} \in \mathbb{R}^{M\times 3}$ is defined as:

$$\mathbf{N} = \begin{bmatrix} n_{1_x} & n_{1_y} & n_{1_z} \\ \vdots & \vdots & \vdots \\ n_{M_x} & n_{M_y} & n_{M_z} \end{bmatrix}. \tag{3}$$

One could argue that plane matching is a variant of data association. The data association step of ICP typically involves a local point search within each pair of patches. Instead, we consider simple geometric models (the bounding boxes) to match the plane patches directly, instead of their points. This global search is easier to solve and does not depend on local density of the points. This approach will be further explored in future work by considering convex hulls as in [11].

Computing Alignability: We consider the constraints imposed on the pose of the sensor by the current measurement with respect to a previously captured point cloud. Using a formulation similar to [10], we consider two point clouds captured sequentially in time and the constraint between a current measurement point $\mathbf{p}_j$ and a measurement point $\mathbf{p}_{i,t} \in \mathbb{R}^3$ at time t in the past, both lying on the same plane:

$$\mathbf{n}_p^T(\mathbf{p}_j - \mathbf{p}_{i,t}) = 0 \tag{4}$$

where $\mathbf{n}_p$ is the plane normal. Given the robot position $\mathbf{x} \in \mathbb{R}^3$, we can formulate a second constraint:

$$\mathbf{x} + \mathbf{r}_j = \mathbf{p}_j \tag{5}$$

where $\mathbf{r}_j$ is the ray vector from the current measurement. Substituting Eq. (5) into Eq. (4) and combining the constraints imposed by all measurements in a sweep, we obtain the system of equations

$$\mathbf{N}\mathbf{x} = \mathbf{c}_j \tag{6}$$

where $\mathbf{c}_j$ is a constant vector $[c_1 \dots c_k]^T$. The matrix $\mathbf{N}$ represents the constraints which exist between the reference and reading cloud from the current sensor pose. We can identify the unconstrained dimensions of the system in Eq. (6) by

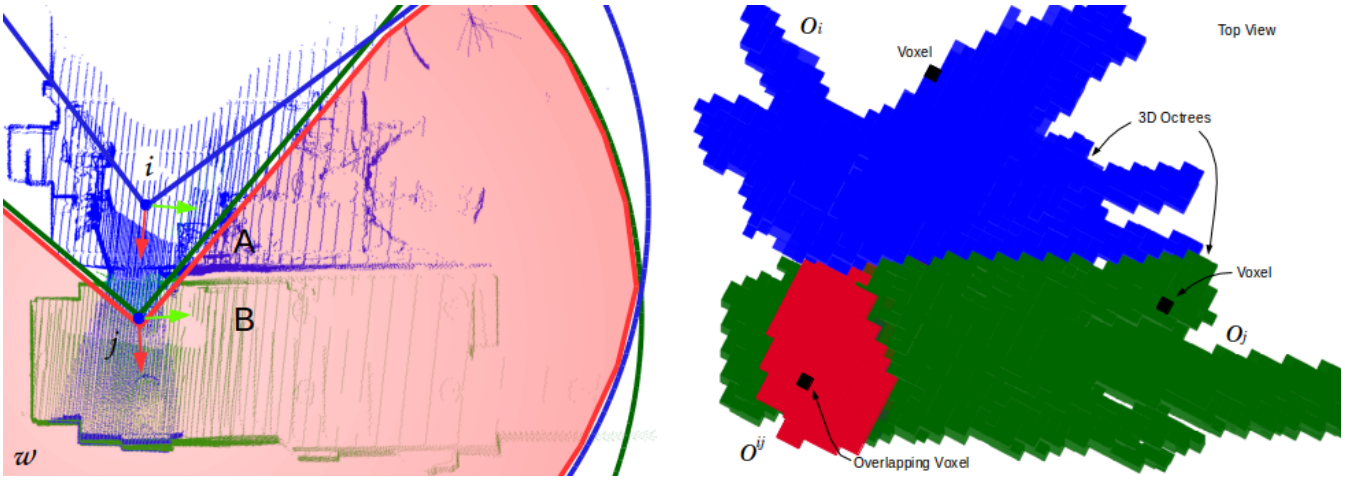

Fig. 4: *Left*: FOV-based overlap parameter [12]. A reference (blue) and reading (green) cloud have been captured from the sensor poses i and j, the latter after crossing a doorway. The region of overlap (red) is delimited by the sensor footprint. In this case the wall between room A and B occludes the view between i and j and the region of overlap is over-estimated. *Right*: The proposed octree-based overlap parameter corrects for the issues occurring in previous work, particularly when occlusions occur. Two octree structures are constructed to model free and occupied space from the two clouds. The volume shown in red, identified by our parameter, reflects what we intuitively think of as volume of overlap.

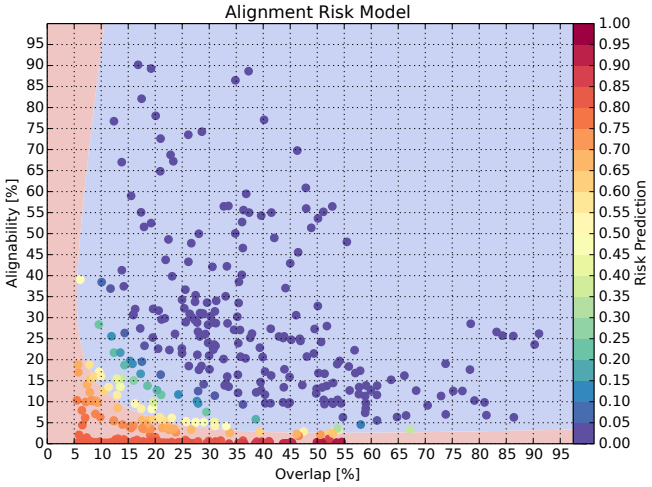

Fig. 5: Alignment risk model learned from overlap and alignability estimates. The classifier has been trained on a set of 1200 binary labelled samples (1:failure/0:success). We show the predictions on the test set, where high risk of alignment failure (red) is expected for low overlap and alignability values, following a polynomial relationship. We observe that using only one parameter (overlap or alignability) is not sufficient, e.g., just using a threshold at 5% on alignability would still accept all samples with overlap below 30%, which would risk a faulty alignment.

performing PCA on the row vectors of $\mathbf{N}$. The structure tensor is of rank 3, hence it follows that the eigenvalues, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$. In Eq. (7) we define *alignability* as the scattering of $\mathbf{N}$:

$$\alpha = \frac{\lambda_3}{\lambda_1}. \tag{7}$$

The scattering parameter is defined as the probability of a set of points to be labelled as a 3D structure [16]. As $\alpha \in [0, 1]$, low values indicate poorly constrained corresponding eigenvector in one of the dimensions.

C. Predicting Alignment Failure

Given these two individual metrics, overlap and alignability, we wish to produce a single combined value which can predict the *risk of alignment failure* $\rho \in [0, 1]$ where a score of zero corresponds to where an alignment will be reliable while a score of one suggests an alignment which is ill-conditioned and will likely fail.

We observed that individually Ω and α cannot predict alignment failure, therefore we define ρ as a function

$$\rho = f(\Omega, \alpha). \tag{8}$$

We learned a model (Fig. 5) for ρ using a third degree polynomial Support Vector Classifier (SVC), which we found empirically to capture the function. Particularly, the prediction ρ is evaluated as the distance of a sample from the optimal hyperplane. Depending on the navigation application, the choice of ρ allows for operation at a preferred point on a Receiver Operator Characteristic (ROC) curve.

The classifier is trained on 1200 real-data observations, as described in Sec. IV. Each observation is registered using AICP and each alignment is manually labelled as a success or a failure.

IV. Experimental Evaluation

So as to validate the study in this paper we carried out a series of experiments using the datasets in Tab. I, as well as a fourth dataset similar to Forum (IF) to train the classifier. The proposed metrics are compared with two baseline terms:

- *Inverse Condition Number (ICN):* the Condition Number is used to determine whether a linear system is ill-conditioned [17]. For the linearized system $\arg\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|$, it is computed as the ratio of the minimum and maximum eigenvalues of $\mathbf{A}^T\mathbf{A}$. For comparison purposes, we consider its inverse and denote it as *ICN*.
- *Degeneracy (D):* for the linearized system above, the degeneracy factor measures the stiffness of the solution w.r.t. disturbances to the constraints [5]. It is computed from the eigenvalues of $\mathbf{A}^T\mathbf{A}$ as $D = \lambda_{min} + 1$.

Our evaluation consists of four experiments:

	Stairs – ST	Apartment – AP	Forum – IF
Source	Pomerleau *et al.* [18]	Pomerleau *et al.* [18]	Ours
Situation	Indoors/Outdoors	Indoors	Indoors
Environment	Structured	Structured, variable scanned volumes	Structured, variable scanned volumes, large occlusions
Dynamics	None	Furniture displaced	Few moving people
Sensor FOV	$\sim 270° \times 270°$	$\sim 270° \times 270°$	$\sim 200° \times 200°$
Path	3D (up a staircase)	2D, small loops	2D, large loop
Scene Area	$(21 \times 111 \times 27)$ m	$(17 \times 10 \times 3)$ m	$\sim (32 \times 27 \times 20)$ m
Scene Sections	1 corridor, 1 staircase, 1 outdoor	5 rooms, 1 corridor	2 rooms, 2 atria, 1 corridor
# of Scans	31	45	117
# of Points	$\sim$ 191000/scan	$\sim$ 365000/scan	$\sim$ 60000/scan
Ground Truth	✓	✓	✗
Application	Exp. C	Exp. C	Exp. D

TABLE I: Features of the datasets used for our experiments.

		True Condition: Alignment Fails	True Condition: Alignment Succeeds
Predicted Condition	Failure	True Positive (TP)	False Positive (FP)
	Success	False Negative (FN)	True Negative (TN)

TABLE II: Confusion map used for our experiments.

Thresholds: *ICN*	*D*	*AR*	
≤ 0.03	≤ 0.06	≥ 0.50	**ST**
≤ 0.05	≤ 0.03		**AP**
–	≤ 0.05		**IF**

TABLE III: Fine-tuned (*ICN*, *D*) and learned (*AR* – ours) thresholds for each dataset.

A) An example comparison between the proposed point cloud overlap parameter Ω and the parameter proposed in our prior work [12]. We demonstrate that Ω is robust in the presence of occlusions.

B) A validation of the *alignability* factor with simulated data. We show that in our experiments *alignability* can predict geometric instability more reliably than *degeneracy*.

C) An evaluation of how the proposed measure of *alignment risk (AR)* outperforms *ICN* and *D* using two standardised datasets from [18]. The experiments created 2986 point cloud alignments and demonstrate the accuracy of our solution with respect to overlap variations.

D) A demonstration of the performance of our localization system on a third dataset, where the proposed method is essential when navigating along corridors and through constrictions and environmental clutter. The system is successful when travelling along a $\sim$ 180 m path. It prevents a total of 21 failures and allows the robot to return to the starting location with a final position error of 0.41% of the trajectory length.

In all our experiments we utilised the thresholds shown in Tab. III when predicting registration failure. We define a failure as 3D pose translational error greater than 0.02 m or rotational error greater than 1°.

A video to accompany this paper is available online.

A. Example Illustrating Overlap

To demonstrate how the proposed octree-based formulation of point cloud overlap better suits real-world scenarios we compared it to our previous FOV-based overlap parameter [12], which was based only on the shape of the sensor's footprint. We now define overlap based on the actual structure of the data rather than the simple sensor's FOV, as discussed in Sec. III-A. A graphical comparison between the two metrics is shown in Fig. 4.

Consider the situation depicted in Fig. 6 (top). During exploration of an indoor environment the robot moves from one room to another through a doorway. At the start, the robot captures a point cloud (shown in blue) which will be used as the reference cloud. Thereafter the robot moves further away from the starting pose (Cloud 2) and the degree of overlap between the reference and the subsequent point clouds decreases – particularly after passing through a doorway.

The plot in Fig. 6 (bottom) shows that the FOV-based parameter over-estimates what we intutively think of as overlap. As the robot enters the new room and turns to the left (Cloud 6), the overlap is clearly low. In contrast, the proposed octree-based parameter successfully identifies this decrease in overlap. Our subsequent experiments will solely use the octree-based overlap parameter.

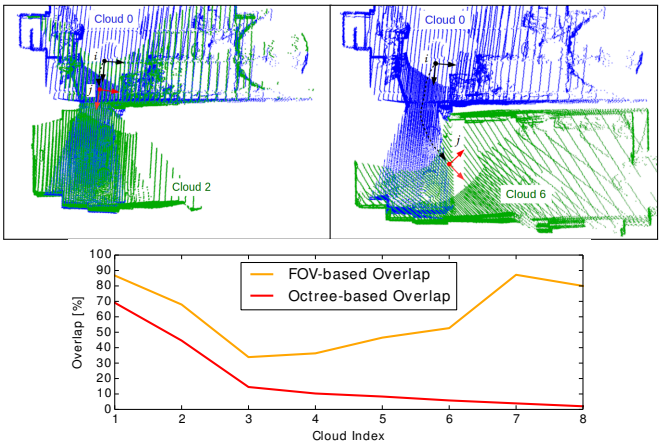

Fig. 6: Exp. A. An example where the basic FOV-based overlap is overconfident while the proposed octree-based parameter correctly estimates overlap. *Top*: Top-down view of the scenario in Fig. 1. A reference and a reading point cloud (blue and green) are captured from the first pose (black) and the current one (red), after crossing a doorway. *Bottom*: Estimate of overlap between the reference and a series of subsequent clouds. Due to its formulation, the FOV-based overlap is overconfident in such a scenario.

B. Validating the Alignability Metric

In a second experiment we aim to show that the proposed *alignability* metric can identify geometric instability in the input point clouds.

We simulate reference and reading point clouds from within a simple cube. We progressively remove surfaces from the cubic reading cloud until the alignment problem becomes unconstrained. We consider the following events, including four constrained (1-4) and five unconstrained geometries (5-9) of the reading cloud: 1) cube, 2) cube missing one surface, 3) cube missing two adjacent surfaces, 4) cube missing three adjacent surfaces and one common corner, 5) cube missing two opposite surfaces, 6) cube missing three surfaces (two opposite), 7) two opposite surfaces, 8) two adjacent surfaces, 9) one surface.

We compare the *degeneracy* and *alignability* parameters using this simulated dataset. For each event the task is to identify geometric instability between the reference and reading cloud. The task was run 100 times per event, with random initial mis-alignment sampled from a zero-mean Gaussian distribution with 0.10 m and 10° variance. The threshold for the two parameters is set to 0.06 and 10 respectively,

and values below threshold indicate unconstrained geometry. Fig. 7 illustrates the result of the experiment. The average accuracy[3] of the two parameters was 79.8% for *degeneracy* and 97.7% for *alignability*. The *degeneracy* factor is computed after the point-wise data association step of ICP, which suffers from higher sensitivity to the initial perturbation. In the case of *alignability* the estimates have lower variance and the results show better separability between constrained and unconstrained events. In our approach we analyse semantic representations (the matched planes), rather than points. This makes it less sensitive to the initial alignment error and more stable overall.

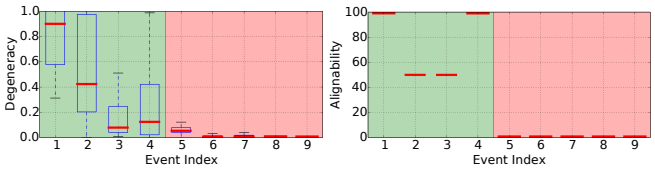

Fig. 7: Alignability validation on the simulated problem in Exp. B. The green and red backgrounds show constrained and unconstrained geometries respectively. The boxplots show the quantiles 5% and 95% (bottom and top end of dashed lines), 25% and 75% (lower and higher end of blue rectangles), 50% (red bars) of the distributions. Note that the boxes for *alignability* are very narrow due to low variance.

C. Performance on Variable Overlap

In Sec. III we discussed the importance of accounting for variations in spatial overlap when predicting point cloud alignment failures in real-world scenarios. Here we evaluate the sensitivity of failure predictions to variations in the overlap between input clouds. We provide a comparison between our *alignment risk (AR)* against *ICN* and *D*.

In this experiment we use two publicly available datasets – Stairs (ST) and Apartment (AP), which consist of a Hokuyo UTM-30LX-EW planar laser mounted on a tilting unit. Ground-truth is available using an external tracking system.

For each dataset we perform a pair-wise alignment between the point clouds. We initialized the reading clouds with random perturbations sampled from a zero-mean Gaussian distribution with 0.10 m and 10° variance. We compute the overlap and measure *ICN*, *D*, *AR* for each alignment. In order to predict a registration failure, we selected a fixed threshold for AR, which is the optimal one learned by our model. In the case of *ICN* and *D* a threshold had to be fine-tuned specifically for each dataset, as detailed in Tab. III. The results are presented in Fig. 9.

Fig. 9a,e illustrate the matrix of overlap estimates using our proposed Ω parameter for both datasets. The diagonal elements show high overlap between each cloud with itself. Note that our estimates are symmetric by formulation.

Fig. 9b-d and Fig. 9f-h present the confusion matrices for *ICN*, *D* and *AR*, which are marked as illustrated in Tab. II.

Considering a localization task, we prefer predicting there will be a failure where there is none (false positive – **FP**) rather than the opposite (false negative – **FN**). In the regions where overlap is low *AR* predicts fewer false negatives than

[3]Accuracy is given by the number of true positives and true negatives by the total population.

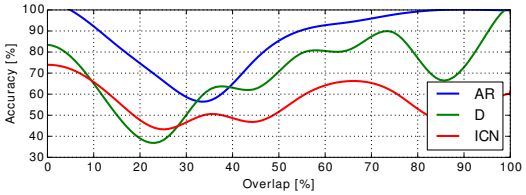

Fig. 8: Exp. C. Accuracy of failure predictions with respect to overlap variation.

ICN and *D*. For *ICN* and *D* the number of point matches is insufficient for the eigenvectors to stabilize. Our approach which measures alignability and overlap, results in a robust prediction.

ICN and *D* often predict false positives along the diagonal despite high overlap. This suggests sensitivity of the parameters to initial perturbations. On the other hand, *AR* is accurate both for low and high overlap.

Fig. 8 shows the accuracy of failure prediction as a function of overlap for each of the three parameters. The results indicate that *AR* has higher accuracy overall. As the overlap reaches 30-40% we note that the success/failure of the alignment becomes less predictable. Nevertheless, our approach performs competitively.

D. Online Performance Analysis

AICP is used as the localization framework for this last experiment, because of the increased robustness to initial alignment error and variation in overlap. The approach leverages a low drift proprioceptive-based state estimate (for example we used the wheel odometry information) to initialize the alignment. In [19], we extended the AICP algorithm to trigger a reference point cloud update when overlap decreases below an empirical threshold. In contrast, in the system presented here, we aim for continuously reliable localization by applying a registration correction to the robot state estimate depending on the predicted risk of alignment failure. When the risk is high, we query a reference point cloud update and rely on proprioception until the next laser measurement becomes available. Furthermore, we replace the original FOV-based overlap parameter with our octree-based one.

We test our localization system on the IF dataset. Our dataset is collected by a Clearpath Husky mobile robot equipped with a Carnegie Robotics Multisense SL. This sensor is composed of a stereo camera and a Hokuyo UTM-30LX-EW planar laser spinning about the forward-facing axis. Every few seconds it spins half a revolution and a 3D point cloud is accumulated. The speed of rotation of the device is set to 15RPM as a compromise between density of the clouds and accumulation time.

For this experiment we can align in three dimensions (x, y and yaw) and use wheel odometry to estimate roll, pitch and z. We predict registration failures using either *D* or *AR*, for comparison purposes. The robot navigates along a $\sim$ 180 m path while exploring the area shown in Fig. 10, which includes two cluttered rooms, two wide atria with high ceilings and a corridor. The exploration involves passing across constrictions such as doorways.

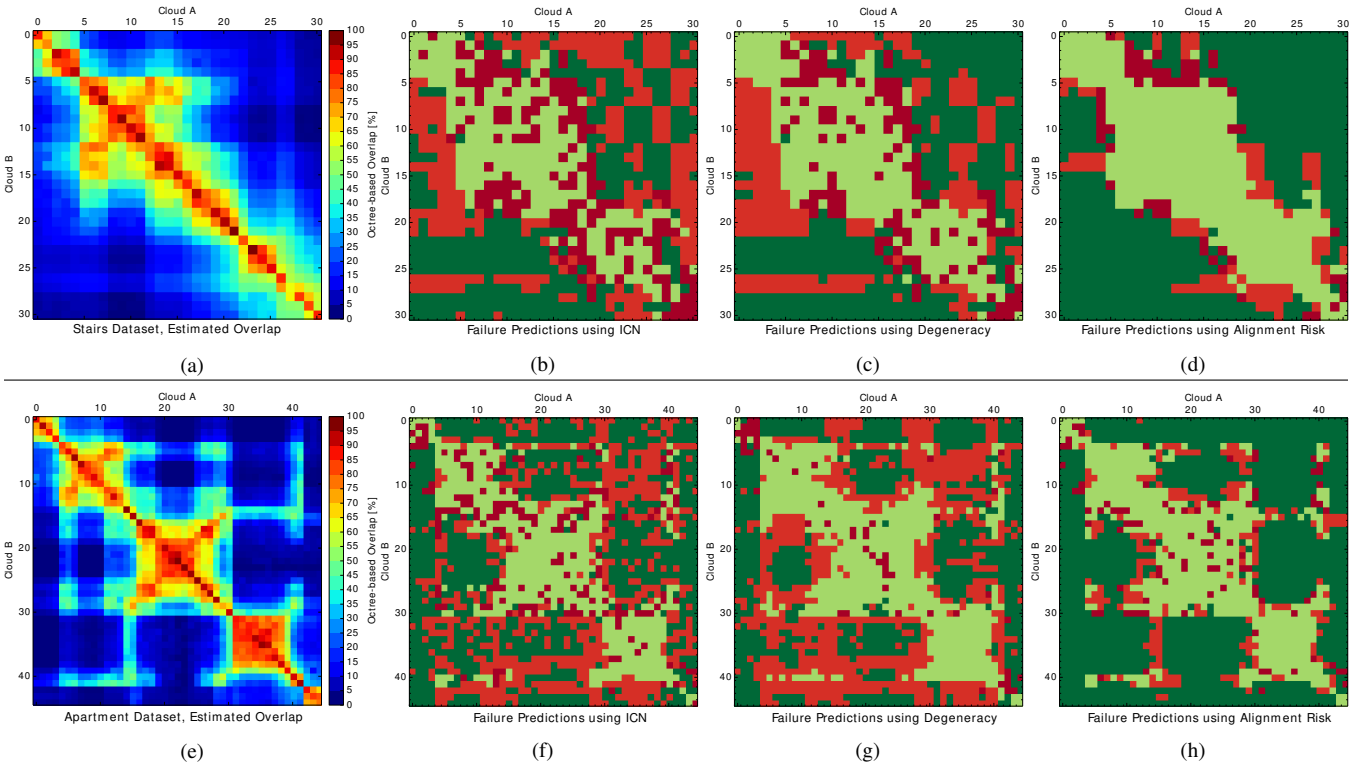

Fig. 9: Exp. C. *Left:* Estimated octree-based overlap for ST and AP datasets. Arrays can be read as the volume of common space between Cloud A and B. High overlap is marked in dark red, while low overlap in dark blue. *From right to left*: failure prediction results for *AR*, *D*, *ICN* on the ST (*top*) and AP (*bottom*) datasets. The matrices follow the color-coding convention in Tab. II, i.e., **TP**, **TN**, **FP**, **FN**.

The localization results, qualitatively evaluated from careful observation of the map built without loop closures in Fig. 10, are such that the estimated trajectory is close to error free when using *AR*. By aligning the last point cloud (red) into the first one (black) we estimate a pose error at the end of the run of 0.73 m in translation, corresponding to 0.41% of the trajectory length, and 0.71° in rotation. This result demonstrates improved localization reliability when using *AR* for failure prediction. The parameter could predict and prevent all failures during this experiment, while being robust to geometric instability and overlap variation occurring in a real-world scenario.

In Fig. 11 we show the registration failures identified by *D* (top plot) and *AR* (bottom plot). *D* could not detect the alignment failures at locations 1, 2, 3, 4 which caused the system to lose track of the robot's trajectory. As shown in the images in Fig. 11 (right) the mis-predictions occurred in situations where the overlap between the input clouds was low, with a consequent instability of the *D* factor. Due to these missed detections we could not complete the run on this dataset when using *D* for failure prediction.

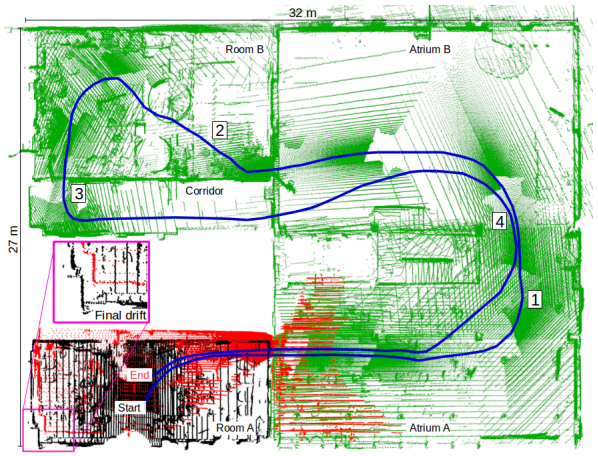

Fig. 10: Top-view of the map reconstructed during Exp. D. No loop closures have been performed. The trajectory estimated by our localization system using the proposed *AR* for failure prediction is depicted in blue. Along the path *AR* captured all 21 alignment failures. We indicate with numbers the locations where only *AR* was successful (constrictions – 2, 3, atria with occlusions – 1, 4), with correspondence in Fig. 11. The pose error at the end is computed from the transform which aligns the last cloud (red) into the first one (black) and results in 0.41% of the $\sim$ 180 m trajectory length.

V. Discussion

The proposed approach is able to reliably prevent localization failures in real-world scenarios where geometric instability and overlap variations occur frequently and challenge point cloud registration algorithms.

In our work we recognized some fundamental advantages in using a learning based approach: i) during all experiments on different datasets we selected a fixed threshold for *AR*. This threshold is the optimal one learned by our model. ii) depending on the application, we might want to be very robust to false negatives, preventing localization failures and hard recovery. The trade off for this scenario comes at the cost of accepting more false positives and reference cloud updates. In turn, each reference cloud update introduces a small incremental error to our estimate. This is useful when a low drift odometry prior is available, allowing us to

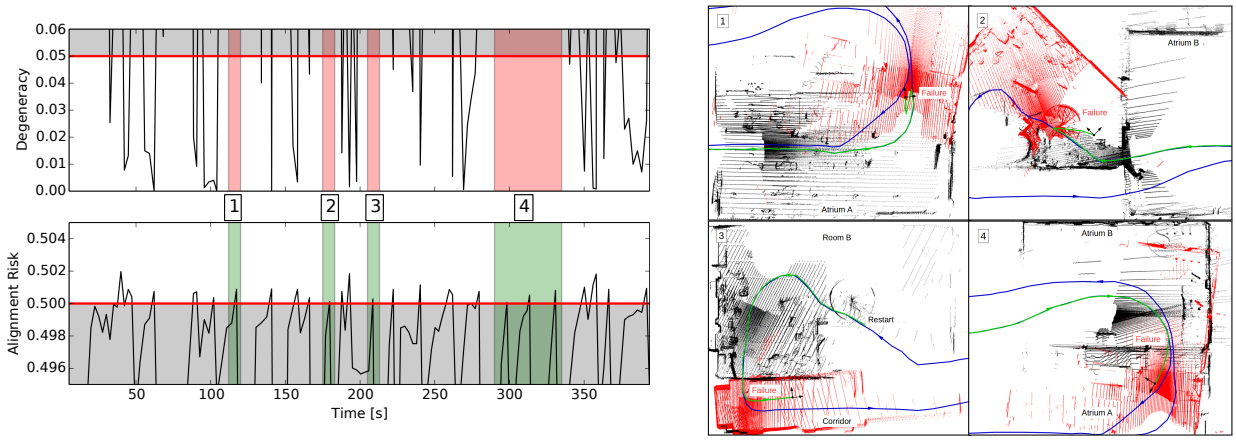

Fig. 11: *Left:* Failure predictions during Exp. D. During navigation D could not detect 4 alignment failures, as highlighted by the red sections on the top plot. We indicate each event with a number (1, 2, 3, 4). A top view of these events is shown on the right. The trajectory estimated using D for failure prediction is depicted in green, whereas the one estimated using the proposed AR is blue. AR successfully predicted all 21 failures.

lower the AR threshold to accommodate more false positive predictions.

Additionally, we believe that our model could facilitate the search of loop closures in a SLAM system, by sorting the candidates by AR score prior to data association.

VI. Conclusions

In this work we proposed a strategy for point cloud alignment failure prediction. We explored the degree to which alignment failure is affected by geometric instability of the input point clouds, as well as the spatial overlap between them. We adopted a data-driven approach to evaluate the geometric constraints available for alignment and the volume of spatial overlap between the clouds. We used this data to learn a model to predict the risk of a failed alignment.

This allows us to be independent of the adopted registration strategy and point-wise data association, as well as to easily select the optimal threshold learned by our model in order to predict a failure, which avoids manual fine-tuning.

We evaluated our approach on different datasets and provided comparisons to existing techniques. Our algorithm overcomes the weaknesses of the baseline techniques identified in the context of real-world scenarios, where constrictions and occlusions cause reduced overlap between observations. We demonstrated how our approach can help improve the reliability of laser-based localization systems during exploration of unknown and cluttered man-made environments. In a large indoor exploration demo the system was able to reliably estimate the robot state with a final pose error of 0.41% of the trajectory length, and to build an accurate 3D representation of the environment.

Future work will focus on the extension of this approach to a SLAM system with loop closure detection.

References

[1] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Trans. Robotics*, vol. 32, no. 6, pp. 1309–1332, Dec. 2016.

[2] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing ICP variants on real-world data sets," *Autonomous Robots*, vol. 34, no. 3, pp. 133–148, Apr. 2013.

[3] J. Guehring, "Reliable 3D surface acquisition, registration and validation using statistical error models," in *Intl. Conf. on 3-D Digital Imaging and Modeling (3DIM)*, May 2001, pp. 224–231.

[4] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy, "Geometrically stable sampling for the ICP algorithm," in *Intl. Conf. on 3-D Digital Imaging and Modeling (3DIM)*, Oct 2003, pp. 260–267.

[5] J. Zhang, M. Kaess, and S. Singh, "On degeneracy of optimization-based state estimation problems," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, May 2016, pp. 809–816.

[6] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.

[7] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image and Vision Computing*, vol. 10, no. 3, pp. 145–155, Apr. 1992.

[8] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Robotics: Science and Systems (RSS)*, June 2009.

[9] J. Serafin and G. Grisetti, "NICP: Dense normal based point cloud registration," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, Sept 2015, pp. 742–749.

[10] W. Zhen, S. Zeng, and S. Soberer, "Robust localization and localizability estimation with a rotating laser scanner," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, May 2017, pp. 6240–6245.

[11] K. Pathak, D. Borrmann, J. Elseberg, N. Vaskevicius, A. Birk, and A. Nüchter, "Evaluation of the robustness of planar-patches based 3D-registration using marker-based ground-truth in an outdoor urban scenario," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, Oct 2010, pp. 5725–5730.

[12] S. Nobili, R. Scona, M. Caravagna, and M. Fallon, "Overlap-based ICP tuning for robust localization of a humanoid robot," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, May 2017, pp. 4721–4728.

[13] M. Pfingsthorn and A. Birk, "Generalized graph SLAM," *Intl. J. of Robotics Research*, vol. 35, no. 6, pp. 601–630, May 2016.

[14] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, pp. 189–206, Apr. 2013.

[15] T. Rabbani, F. van den Heuvel, and G. Vosselman, "Segmentation of point clouds using smoothness constraints," *ISPRS Annals of the Phot., Remote Sens. and Spatial Inf. Sciences*, vol. 35, pp. 248–253, 2006.

[16] M. Weinmann, B. Jutzi, and C. Mallet, "Semantic 3D scene interpretation: A framework combining optimal neighborhood size selection with relevant features," *ISPRS Annals of the Phot., Remote Sens. and Spatial Inf. Sciences*, vol. 2, no. 3, pp. 181–188, 2014.

[17] E. W. Cheney and D. R. Kincaid, *Numerical Mathematics and Computing*, 6th ed., 2007.

[18] F. Pomerleau, M. Liu, F. Colas, and R. Siegwart, "Challenging data sets for point cloud registration algorithms," *Intl. J. of Robotics Research*, vol. 31, no. 14, pp. 1705–1711, Dec. 2012.

[19] S. Nobili, M. Camurri, V. Barasuol, M. Focchi, D. Caldwell, C. Semini, and M. Fallon, "Heterogeneous sensor fusion for accurate state estimation of dynamic legged robots," in *Robotics: Science and Systems (RSS)*, July 2017.